VOL. 12, NO. 13, JULY 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

CHARGER MONITORING STREAM ANALYSIS USING ARTIFICIAL NEURAL NETWORKS

Junghoon Lee and Gyung-Leen Park
Department of Computer Science and Statistics, Jeju National University, Republic of Korea
E-Mail:jhlee@jejunu.ac.kr

ABSTRACT

This paper develops a trace and prediction model for the energy load imposed by electric vehicle charging in Jeju City, based on artificial neural networks as well as the massive amount of charger monitoring streams collected for about 1 year. The monitoring system generates a data archive, in which a single report essentially embraces charger ID, timestamp, and battery level reading. Our data analysis module, consisting of MySQL, R, and C program implementation across Linux and Windows PC domains, not only processes raw data files to insert into a database table but also finds the active charger set. The distributions of the number of reports and the power consumption for each charger are investigated. For the day-by-day power consumption stream for vehicle charging, a neural network model traces and predicts the time series having 124 sequential values. The experiment result shows that our model catches up with unpredictable spikes just with small time lag.

Keywords: electric vehicle, charger monitoring, big data analysis, artificial neural network, load prediction model.

1. INTRODUCTION

The penetration of EVs (Electric Vehicles) is empowered by the expansion of battery capacity and the construction of charging facilities [1]. Many cities are trying to provide public chargers, be it free or not, to EV drivers over the wide area to overcome long charging time and short driving range, compared with the gasoline-powered vehicles. The chargers or charging stations are generally managed by an administration authority, possibly taking advantage of information and communication technologies, to support stable and safe operations of a charging service [2]. Essentially, the manager continuously monitors and stores the current state of each charger to check whether it is operating without any problem and to collect the amount of electricity ejected to EVs. This process inevitably creates a massive amount of monitoring data, just like other smart grid entities [3]. The availability of such big data can make it possible to plan an additional facility construction in the target city, develop a new business model, and the like [4].

As one of the most prominent cities leading a variety of smart grid efforts, Jeju City, Rep. of Korea, is installing more and more chargers, standardizing the message format between chargers and management authorities, and testing the city-level data concentration system for EV charging. During the test period lasting about 4.5 months, 50 chargers have generated quite stable reports, each of which includes timestamp, charger ID, and amount of charged energy from the beginning of the charger operation. Even though some fields, for example, the plug-in time of an individual EV and the exact location of each charger are hidden for personal information protection, this system gives us a lot of valuable information, especially benefitting from the time-series analysis techniques [5]. Hence, this paper introduces the features of our charger monitoring system, describes the data processing architecture, and analyzes the monitoring stream mainly with the ANN (Artificial Neural Network), which is one of the most widely used for nonlinear series modelling [6].

The rest of this paper is organized as follows: After issuing the problem in Section 1, Section 2 reviews some related work and research background. Section 3 describes the data processing architecture and Section 4 analyzes the statistics obtained by the charging data stream. Finally, Section 5 concludes this paper and introduces future work.

2. RELATED WORK

Basically, EV chargers convert AC power to DC power, playing a role of an electric load. Power quality management during the period of intensive charging is very important. [7] points out that power quality metrics include voltage fluctuation, power factor, 3-phase unbalance, and the like. This approach takes advantage of monitoring data streams acquired by probing at the 380V bus in an EV charging station. The main focus is put on harmonic current emissions and the power factor of 4 types of chargers, namely, on-board slow charger, off-board fast charger, and two multi-mode off-board chargers. A charging station having tens of chargers sharing a common bus usually suffers from harmonic currents, mainly due to the difference in not only charging power amounts but also harmonic phases. After all, the analysis finds out that harmonic voltage deviation falls within the permissible range in their standard, except the case of a single-phase on-board charger.

[2] analyzes the power balance and the impact of EV charging, exploiting the data set acquired from its chargers developed by the University of Beira Interior. The power quality includes steady-state power quality variations and momentary disturbances that may impact the system load. The charging profile is obtained by Fluke 434 Series II, while supplying electricity to an EV, specifically, a Renault ZOE model having 22 *kwh* lithium-ion battery capacity. The main measurement parameters include the

voltage, current, and active-reactive power. The experiment observes drastic harmonic load jumps which may lead to unacceptable voltage deviations as well as harmonic power losses. Anyway, according to their experiments, EV fast charging can be a good power factor compensator in big industries.As such, many approaches are built on top of a data monitoring system, while ours is challenging a citywide charging infrastructure.

As an effort of big data analysis on smart grid domains, our research team has built a data processing framework for EV battery consumption on Jeju Island [8]. The SoC (State of Charge) changes along the main roads in this area are captured and stored in separate files for each trip. The analysis module first supports diverse queries to find a specific set of records, traversing the directory hierarchy of log files. Such queries can locate high-speed intervals, the number of records belonging to a target region, and the overall speed across a road segment. In addition, the movement tracker module calculates the driving distance from the start point to obtain distance-SoC curves. Here, each location stamp is given in the WGS84 coordinate. The inter-record correlation allows us to find the road interval in which the SoC level drops very sharply, so installation of fast chargers is required. The framework integrates the road network given byshape files and the retrieval results are displayed on the map. Here, we can alter the weight of each link with such factors as battery consumption.

3. DATA PROCESSING

To begin with, the real-time monitoring system displays the current state of each charger via a geographic information-based application. The archive of such monitoring stream is given to our analysis module as shown in Figure-1 on as-wanted basis. Our raw data processor, implemented in the C language, converts the log files into a series of SQL *insert* statements, each of which adds a status report into the predefined database table. Original log files include history ID, user ID, status code, timestamp, and power consumption. However, not every record has all fields; some are missing, according to chargers and time intervals. As we are mainly concerned with the power consumption, the timestamp and the SoC level of each record are extracted first by a little bit complex string processing. The SQL query file is transferred to a Linux PC via the FTP utility. The SQL statements are executed in a batch mode.

Applications running on Windows PCs can retrieve those table entries stored in Linux MySQL, via ODBC, RMySQL, and FTP, according to the application type. Basically, retrieval queries can be executed on the MySQL command line interpreter, and the result can be downloaded via FTP as a text file. In this case, GNUPLOT can take this file and generate an appropriate graph. Next, with ODBC library integration, it is possible to issue SQL queries in a C language program. After retrieval, we can convert the result to a set of learning patterns to build a neural network model possibly employing open software modules such as FANN (Fast ANN) [9], which provides a diverse ANN application programming interfaces through comprehensive libraries. Moreover, the RMySQL extension allows us to retrieve information using SQL queries even in the R workspace, one of the famous statistics packages nowadays. It is possible to apply a variety of sophisticated data mining schemes, such as fuzzy logic, decision tree, and the like [10].

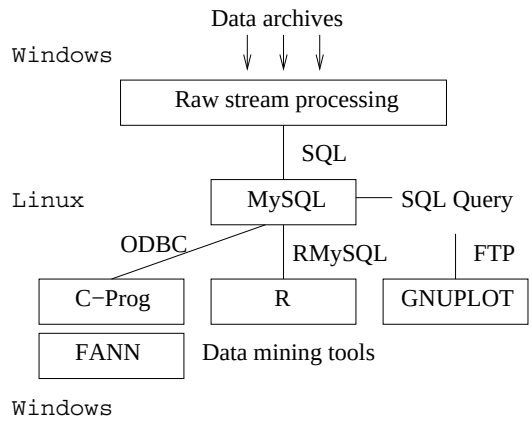

Figure-1. Data processing framework.

3. ANALYSIS RESULTS

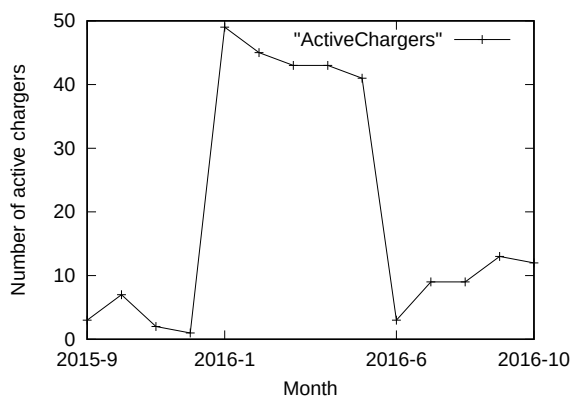

Figure-2. Daily number of active chargers.

The data archive is collected during the period from 2015-9 to 2016-10. However, the system has experienced extensive tests until 2015-12 and been reconfigured in 2016-10, specifically, separation of management domains. Between them, we call it *hot period*. As can be seen in Figure-2, the number of active chargers, which have generated at least one report, reaches 50. However, 12 chargers have been separated from the monitoring system due to some reasons such as device failure and ownership change. After removing the chargers which generated less than 2,000 records, 39 chargers are observed to be meaningful. Those chargers which have generated reports but spent almost no energy are not excluded. During the whole period, 54 chargers have generated 1,691,619 records, while during the hot period, the number of records from the active chargers is

1,478,661. The hot period dominates the whole data set, as shown in Figure-3.

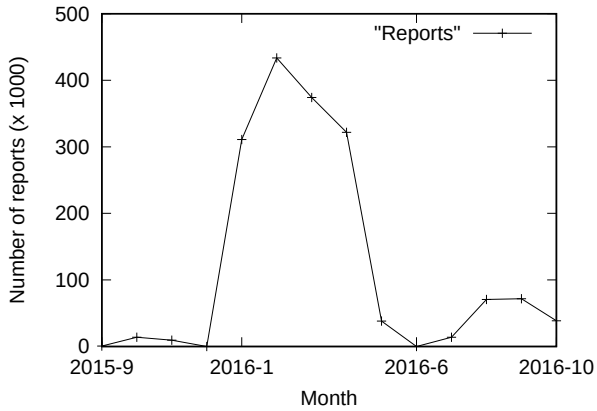

Figure-3. Daily number of valid records.

Next experiment looks into the distribution of the number of records and the power consumption for each active charger during the hot period. In Figure-4 and Figure-5, chargers are sorted by the number of reports and the accumulated consumption amount, respectively. Here, the chargers have not been reset during the investigation interval and the reported amount increases monolithically. Hence, the sequences in both figures do not coincide and the x-axis does not indicate actual charger numbers. As shown in Figure-4, the number of records for each charger ranges from 21,475 to 45,153. This difference comes from message loss, unequal operation time length, and the operation mode change. Basically, the report interval is 5 minutes. During the sleep mode, the report interval is extended to 10 minutes. The figure shows that the number of reports is evenly distributed for the active charger set, and sufficiently large to discover a specific utilization pattern among chargers.

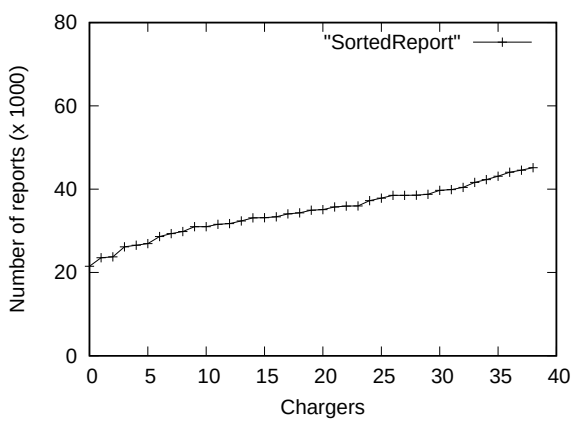

Figure-4. Per-charger number of reports

In addition, Figure-5 plots the power consumption for each charger. It is calculated by the difference between the maximum and the minimum of accumulated power amount. As contrast to Figure 4, the difference between chargers is very significant. According to our inspection, 5 chargers have never been used for EV charging during the whole operation period, while 3 chargers are intensively used by many EVs, showing the accumulated consumption of more than 7,000 *kwh*. All of them have reported at least 20,000 times. The records from those frequently used chargers tell that they almost always have EVs plugged-in. definitely, those chargers are located in office buildings possessing many EV memberships. For the top charger which supplied about 9,000 *kwh*, considering that battery capacity is usually around 18 *kwh*, it has served at least 500 times for 5 months. This service ratio means that that charger supplied electricity roughly to 3 EVs to their full battery capacity each day on average.

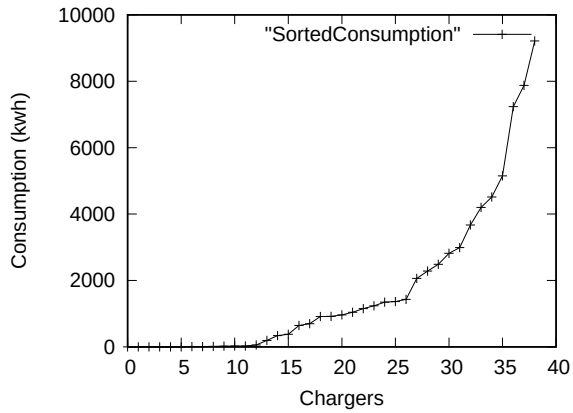

Figure-5. Per-charger power consumption.

For the inspection of charging pattern for a specific charger, we select the most frequently used charger out of those which have consumed more than average electricity. Figure-6 plots the amount of energy used for EV charging on that charger. Here, the x-axis denotes the elapsed hour from the instant the first report is generated, while the y-axis denotes the accumulated power. It is calculated by subtracting the SoC of the first record from that of a current record. For the report stream of the target charger, the experiment keeps track of timestamp fields to catch the change in the hour substring. How sharp the curve increases depends on how frequently the charger is used, as the energy injection ratio is largely constant. As can be seen in the figure, there exist peak intervals and during the rest of them, the consumption hardly changes. Miscellaneously, small numerical fluctuations are detected in the slope, but the curve looks piecewise straight, absorbing their influences.

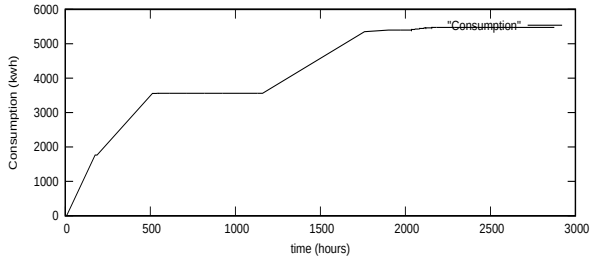

Figure-6. Accumulated consumption for a charger.

Finally, for the sake of measuring the system-wide load imposed by EV charging [7], we retrieve the day-by-day amount of energy consumption for each charger from the database table first. Namely, the whole records are grouped by charger IDs and days. Then, the difference between the maximum and the minimum of power consumption for each day is obtained for per-charger daily consumption. By summing up the charger-by-charger consumption amount for each day, the experiment gets the daily EV charging load in our city. The result is plotted in Figure-7 as a solid curve. The length of the hot interval is 124 days, and the figure shows extremely nonlinear behavior of the EV charging load. It contains unexpected sharp spikes from time to time. We can see that such spikes take place in the latter part of the observation interval more often. As the EV deployment does not reach its saturation point, the charging load may look quite unstable. The charging load is dependent on so many factors including the per-user driving pattern, day-of-week, traffic condition, and the like.

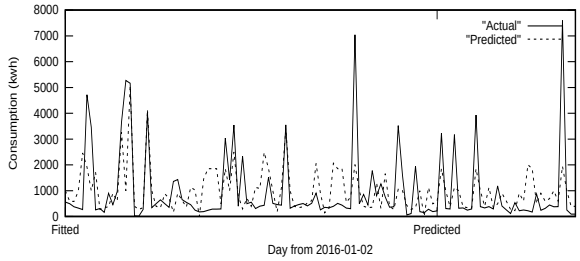

Figure-7. Daily citywide power consumption.

At this stage, we develop a fitting and prediction model for daily EV charging load based on ANN. Specifically, this procedure employs FANN, which is described previously. According to our experience, its convergence is better than others. Moreover, this library is easy to install and can be called from conventional C language programs by linking open object modules. The model is built by making the neural network learn from a set of input-output patterns according to the principle of *learn by example*. For the time series of consumption amounts, the previous 4 elements are taken as inputs while the next one is the sole output. Hence, the number of input nodes and the number of output nodes are 4 and 1, respectively. By trial and error, the number of hidden nodes is set to 15.Out of 124 records; we takethe first 90 for learning pattern generation, while the others are left for the evaluation of prediction accuracy. Then, the learning epochs iterate 30,000 times and finish at the root mean square error of 0.003. In the normalization to the interval of [0.0, 1.0], 1.0 corresponds to the maximum value, namely, 7,593 *kwh*.

The whole set of 124 records are fed to the developed neural network model. As the first 90 are used for the learning phase, the output of the model shows how well it traces the behavior of the given stream. On the contrary, for the rest of them, not used for learning, the result indicates the accuracy of the prediction model. Figure-7 shows these results by the dotted line. Actually, it is not necessary to numerically calculate the error metrics such asroot mean square error, as spikes are very hard to catch up with, extending the average error size. Time lagging also increases the error size, even though the horizontal gap is not so large. According to the figure, during the fitting interval, the model follows not so sharp spikes just with small time lag. In addition, during the prediction interval, namely, the right-hand side of the *Predicted* label in Figure-7, the prediction result tends to proceed less dynamically. In both cases, the biggest spikes are not even closely traced or predicted, as other time series analysis schemes will.

4. CONCLUDING REMARKS

Charging infrastructures are being built in many cities to accelerate the penetration of EVs, and this charging system creates a massive amount of monitoring data not only for real-time management but also for system analysis. In this paper, we have developed a trace and prediction model for the EV charging load in Jeju City, taking advantage of a well-known open ANN library as well as the massive amount of charger monitoring streams. The archive is collected for about 1 year from 54 chargers, consisting of over a million status reports. Each record includes charger ID, timestamp, and amount of consumed energy from the operation start. Our data analysis module, consisting of MySQL, R, and own C program implementation, processes raw data file not only to eliminate the redundancy but also to fill the predefined database table and allow various queries from different application types.

The proposed analysis procedure has identified the active charger set during the stable operation period after investigating the distributions of the number of reports and the amount of electricity used for EV charging for all chargers. To obtain the day-by-day system-wide charging load in our city, a charger-level data cleansing process has been conducted and query programs are carried out. It is shown that the conventional ANN model can trace and predict the EV charging load just with a small time lag in catching up with the highly unpredictable spike patterns. This result shows that the developed data processing can allow us not only to plan charging infrastructure facilitation more efficiently but also to predict the EV charging load according to the increased number of EVs, especially combined with geographic information. This analysis will make is possible to

recommend EVs on where to charge according to diverse options [11].

As future work, we are planning to recognize charger-by-charger energy consumption patterns and group them exploiting the time warping algorithm. It will help us better understand the charging load according to their location. This understanding will allow us to integrate more renewable energy for EV charging to alleviate the energy burden in the system grid and make the EV-based transport more eco-friendly [12]. Moreover, EV battery is expected to be a critical resource for V2G (Vehicle-to-Grid) [13]. It is worth mentioning that Jeju city is making an effort to extend the coverage of solar energy generation.

ACKNOWLEDGEMENT

Prof. Gyung-Leen Park is the corresponding author. This research was supported by Korea Electric Power Corporation through Korea Electrical Engineering & Science Research Institute. (Grant number: R15XA03-62).

REFERENCES

[1] Ipakchi A., Albuyeh F. 2009. Grid of the Future. IEEE Power & Energy Magazine. pp. 52-62.

[2] Pinto R., Pombo J., Calado M., Mariano J. 2015. An Electric Vehicle Charging Station: Monitoring and Analysis of Power Quality. In: 9th International Conference on Compatibility and Power Electronics.

[3] Aman S., Simmhan Y., Prasanna V. 2015. Holistic Measures for Evaluating Prediction Models in Smart Grids. IEEE Transactions on Knowledge and Data Engineering. 27(2): 475-488.

[4] Lee J., Park G., Park C. 2015. Charging Facility Monitoring Stream Analysis based on Hadoop for Smart Grids. International Journal of Software Engineering and Its Applications. 9(11): 153-162.

[5] Liu R., Vellaithurai C., Biswas S., Gamage T., Srivastava A. 2015. Analyzing the Cyber-Physical Impact of Cyber Events on the Power Grid. IEEE Transactions on Smart Grid. 6(5): 2444-2453.

[6] Lee J., Park G. 2014. Design of a Greedy V2G Coordinator Achieving Microgrid-level Load Shift. In: Z. Zeng *et al*. (Eds.) ISNN 2014, 8866: 584-593, Springer, Heidelberg.

[7] Li Q., Tao S., Xiao X., Wen J. 2013. Monitoring and Analysis of Power Quality in Electric Vehicle Charging Stations. In: International Future Energy Electronics Conference.

[8] Lee J., Park G. 2016. Analysis of Stream Data from Electric Vehicles for Energy Consumption Statistics, Advanced Science Letters. 22(11): 3454-3458.

[9] Nissen S. 2005. Neural Network Made Simple. Available at http://fann.sourceforge.net/fann_en.pdf, Software 2.0.

[10] Brunsdon C., Comber L. 2015. An Introduction to R for Spatial Analysis & Mapping. SAGE Publication Ltd.

[11] Qiu G., Wang K., Li S., Dong J., Xie M. 2014. Big Data Technologies in Support of Real Time Capturing and Understanding of Electric Vehicle Customers Dynamics. In: 5th IEEE International Conference on Software Engineering and Service Science. pp. 263-267.

[12] Amato, A. Aversa, R., Martino B., Venticinque S. 2016. A Cyber Physical System of Smart Micro-Grids. In: International Conference on Network-Based Information Systems. pp. 165-172.

[13] Nguyen H., Zhang C., Mahmud M. 2015. Optimal Coordination of G2V and V2G to Support Power Grids with High Penetration of Renewable Energy. IEEE Transactions on Transportation. 1(2): 188-195.